Market participation and sale of potatoes by smallholder farmers in the central highlands of Angola: A Double Hurdle approach

Byron Reyes (reyespad@msu.edu), Cynthia Donovan, Richard Bernsten, and Mywish Maredia
Michigan State University

Selected Poster prepared for presentation at the International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil, 18-24 August, 2012.

Abstract

This paper uses a double hurdle regression analysis to estimate the factors influencing marketing decisions among potato growers in the central highlands of Angola, focusing on gender of household head, productive asset ownership and transaction costs. Although the results suggest that the quantity produced is exogenous in the models for market participation and for quantity sold, the methodology used provides a framework for others to follow when endogeneity is suspected in one or more variables. The wealth analysis suggests that potato growers, potato sellers and male heads were richer than their counterparts. The linear regression results on quantity produced suggest that female-headed households produced less than their male counterparts, owning productive assets or having access to public assets had no statistical effect on production, and that farmers who used fertilizer produced more than farmers who didn't apply fertilizer to their fields. The double hurdle regression results suggest that (1) male-headed households were more likely to sell potatoes, (2) owning productive assets and having access to government extension services, conditional on market participation, positively affected the quantity sold, (3) transaction costs, conditional on market participation, negatively affected the quantity sold, and (4) quantity produced was a marginally significant positive factor on both the likelihood of selling potatoes and the quantity sold. In contrast, the unconditional average partial effects suggest that, (1) potato sales were gender neutral, (2) owning productive assets had no statistical effect on quantity sold, (3) transaction costs negatively affected the quantity sold, and (4) having access to extension services and the quantity produced both positively affected the quantity sold. Thus, to boost sales, investments may be needed to promote farmer participation in organizations and/or establish farmer organizations in villages without them, increase farmers' access to extension services, invest in infrastructure, and help farmers increase their production.

1 Introduction

Agricultural households can be classified into three categories based on their net position relative to the market: net buyers, net sellers, or autarkic (non-participants).[1] It is known that market participation is both a cause and a consequence of development (Boughton *et al*. 2007; Barrett 2008). Markets provide households the opportunity to benefit from trade; i.e. they can sell their surpluses and purchase goods and services they need, according to their comparative advantage (Boughton *et al*. 2007; Barrett 2008). Furthermore, as households' income increases, the demand for goods and services also increase, hence stimulating development (Boughton *et al*. 2007). However, the net position of the households not only depends on market prices; it also depends on households' access to productive technologies (e.g. improved varieties, inputs, etc.) and adequate private and public goods (Barrett 2008) and services.

To date, price-based, top-down macro and trade policy interventions have not been enough to stimulate smallholder market participation and agricultural and rural transformation as expected (Barrett 2008). However, understanding the impact of these policies on smallholder farmers' market participation is important. The fact that market participation is heterogeneous has important implications when studying households' response to governmental policy interventions and should be considered in policy response estimation (Key *et al.* 2000).

It is known that farm households are typically located in environments characterized by a number of market failures[2] (Sadoulet and de Janvry 1995, ch. 6, pg. 9). These authors point that any market could fail for a particular household when the margin between the low price at which the household could sell a commodity and the high price at which it could buy it is large; hence

[1] Goetz (1992) called this classification the household trichotomy.

[2] De Janvry *et al.* (1991) demonstrated that market failure was household, not commodity specific.

the household may be better off by being autarkic. This leads us to the concept of price bands widely described in the literature (De Janvry *et al.* 1991; Sadoulet and de Janvry 1995; Key *et al.* 2000), which refers to the effective price paid by buyers and received by sellers (both market participants) and that determines the household's net market position.

To boost market participation, one of the government's goals should be to make investments targeted at reducing the magnitude of the price band. This magnitude could be affected by transaction costs, the existence of shallow local markets,[3] and price risks and risk aversion (Sadoulet and de Janvry 1995). This paper studies the effect of gender of head, transaction costs and productive asset ownership on household's marketing decisions, using cross-sectional data from three provinces of Angola.

Angola ended its 27-year long civil war in 2002. The war had a large impact on the country's infrastructure[4] and caused the demise of the rural economy and the subsequent sharp rise in the urban areas (World Bank 2007). Other effects were the loss of life of over 1 million people and migration (rural to urban but also to neighboring countries) (World Bank 2007). Migration, urbanization, population growth, and increasing household incomes have caused an increase in the demand for food in the major cities of the country. For example, the estimated 2005 annual demand for potatoes, onions, carrots, and dry beans in Luanda (the capital city) was a little over 197,000 MT, 61% of which was imported from neighboring countries, especially South Africa (World Vision 2008).

[3] For details see Sadoulet and de Janvry (1995, ch. 6, pg. 9). The idea is that there is a negative covariation between household supply and prices because when the harvest is good and surplus could be traded, the price falls because all other households also have good harvests, making the price band to widen (the opposite is also true).

[4] It is estimated that US $4 billion will be necessary just to restore the road and bridge network of the country (World Bank 2007).

Although expenditures in energy, agriculture, mining, and transportation were high (10.2% of GDP; US$1.4 billion) in 2003, by 2005, expenditure in these areas was drastically reduced to only 2.2% of GDP (or US$734 million) (World Bank *et al.* 2007), suggesting that rural households may still face many limitations to actively participate in markets and satisfy part of the demand for food.

In addition to the country's transition from war to peace, the country went (as many other African countries) from a centralized market to a free market (Munslow 1999). However, many food aid programs favor unfair competition and government control has resulted in a poorly developed trading network (World Bank 2007). Furthermore, Angola has been cut off from technological advances (e.g. new varieties) and increasing farmers' productivity[5] still remains a challenge because of the disadvantages of Angola's strong currency and high transportation costs, which discourages competitiveness (World Bank 2007).

The study's main focus is on estimating the determinants of market participation and quantity of potatoes sold, focusing on the effect of gender of the household head, transaction costs and productive asset endowments on marketing behavior, following Boughton *et al*. (2007), Barrett (2008), and Bellemare and Barrett (2006). However, this study implements a double hurdle regression approach and estimates the unconditional (on market participation) average partial effects for the quantity of potatoes sold.

2 Research Gap

Many studies related to the analysis of market participation by agricultural households have focused on (1) dealing with potential problems of sample selection bias when testing

[5] Although Angola enjoys better rainfall than many of its neighbors, crop yields are much lower.

hypotheses about market participation and (2) understanding the role of transaction costs and market failures on households' marketing decisions.

Heckman (1979) discussed sample selection bias as a specification error and provided a technique that allowed for the use of simple regression to estimate behavioral functions free of selection bias in the case of a censored sample. The solution proposed by Heckman (1979) to obtain unbiased estimators was simple.[6] *First*, he demonstrated that the bias that results from using (non-randomly) selected samples could arise from a problem of omitted variables. *Second*, he proposed that, for the full sample (e.g. trainees and non-trainees), a probit analysis could be used to estimate the probability that an individual may be in the selected sample (e.g. will participate in training). *Third*, he demonstrated that by using this probability as a regressor in the equation of interest (e.g. trainees' earnings) one could obtain unbiased estimators.

In its widely cited work, Goetz (1992) modeled the agricultural household's discrete decision of whether to participate in markets separately from the continuous decision of how much to trade, conditional on market participation;[7] an innovation in market participation analysis at the time. Elaborating on the groundbreaking work of Goetz (1992), Key *et al.* (2000) studied the effect of proportional and fixed transaction costs on household supply response. They implicitly modeled the household as making the discrete market participation choice simultaneously with the continuous decision of how much to trade.[8] In constructing their agricultural structural household model, they kept separated the structural supply functions from the production threshold functions. By estimating this model, they were able to separately

[6] See Heckman (1979) for a detailed explanation.

[7] That is, he assumed households make sequential choices: they *first* decide whether to participate in the market; *then*, conditional on participation, they decide how much to trade.

[8] In contrast to Goetz (1992) who assumed households make sequential marketing choices.

identify the effect of proportional and fixed transaction costs on supply response, while avoiding the problem of selection bias described by Heckman (1979).

As noted, some authors assume households make marketing decisions sequentially while others assume they make these decisions simultaneously. Bellemare and Barret (2006) developed a two-stage econometric method that allowed them to test whether rural households in developing countries make market participation and volume decisions simultaneously or sequentially. They found evidence in favor of sequential decision making, with the implication that households that make sequential marketing decisions are more price-responsive and less vulnerable to trader exploitation.

Although many recent studies have focused on the effect of transaction costs, farmers' assets and wealth also affect marketing decisions. Boughton *et al.* (2007) took an asset-based approach to analyze smallholder market participation in Mozambique. They assumed households as making sequential marketing decisions[9] and developed a simple structural model of the household's choice problem, facing two constraints: budget and asset allocation constraints.

Fafchamps and Vargas-Hill (2005) analyzed the factors associated with coffee producers' decision to sell at the market vs. at the farmgate. Although their study didn't focus on the decision to participate in the market,[10] it provides insights about why farmers choose different places for their sales. Barrett (2008) provides a detailed literature review about evidence on smallholder market participation in eastern and southern Africa, focusing in staple food-grains markets.

---

[9] Similar to Goetz (1992).

[10] All coffee producers are sellers because coffee is a cash crop. Therefore, household consumption may be very small, if any.

Markets rarely work perfectly. Household modeling under missing markets is well explained in Sadoulet and de Janvry (1995). De Janvry *et al.* (1991) analyzed the effect of missing markets on farmers' supply response and found that programs directed at reducing the incidence of market failures[11] are very important to increase the supply elasticity of households--hence increasing household's response to price incentives.

The contribution of this study is as follows. *First*, it provides new empirical results to the rather limited literature on market participation in Africa, especially in Angola, by looking at farmers' participation in the potato market in the central highlands of the country. *Second*, besides focusing on transaction costs, this study also focuses on the effect of productive asset ownership[12] and gender of household head on marketing decisions.[13] *Third*, it uses a double hurdle approach to control for self-selection bias and provides unconditional (on market participation) effects of the variables on sale of potatoes.

This paper focuses on potatoes because (1) this crop is very important in the country's agricultural sector because of its high potential to generate profits to smallholder farmers; (2) there is a large unmet demand for potatoes in large cities of the country that currently is satisfied by imports from neighboring countries; and (3) recent private and public investments targeted at improving supply chains in rural Angola are focusing on potatoes (World Vision, 2008). Therefore, generating information about the factors affecting smallholder farmers' marketing decisions will be valuable to target assistance to farmers.

---

[11] For example, infrastructure investments (which reduce transaction costs), better circulation of information on prices, access to credit markets (an indirect source of market failure), etc.

[12] Bellemare and Barrett (2006) did not explicitly study the effect of productive assets on marketing behavior, as Boughton et al. (2007) did.

[13] Recent private investments are interested in learning about the role of gender on household decisions, especially because after the war, many households are lead by females.

3 Research Questions

Although the study's main objective is to generate information about the factors affecting smallholder farmers' marketing decisions, it also attempts to answer the following research questions:

- What are the characteristics of farmers who trade potatoes in the central highlands of Angola, compared to non-traders?
- What factors affect farmers' potato production?
- What factors are influencing farmers' decision of whether to sell their surpluses in the market?
- Conditional on market participation, what factors are affecting the quantity of potatoes traded by farmers?
- What is the unconditional effect of gender of household head, productive asset ownership and transaction costs on the quantity of potatoes sold?
- What policy recommendations could be generated, based on the empirical evidence, to boost market participation?

4 Conceptual Framework

In this section, first, the economic rationale for analyzing household's marketing decisions is explained. Then, an econometric framework is presented to be able to empirically estimate the economic model while addressing the econometric challenges of the analysis.

4.1 Economic model

To analyze the factors associated with farmer's marketing decisions, following Boughton *et al.* 2007 and Barrett 2008, a simple model of household choice is developed. It is assumed that households will maximize their utility *U*, by consuming a vector of agricultural commodities, s^C, for *c* crops, and a Hicksian composite of other tradables, *x*. It earns income from production and possibly sale of any or all crops, and possibly off-farm income, *Y*, which could be earned or unearned.

Crop production is determined by a crop-specific production technology, $f^{C}(A^{C}, G)$, which depends on the flow of inputs (e.g. fertilizer, pesticides, seed, labor) and services provided by privately held quasi-fixed productive assets, represented by the vector ***A***. This function is also affected by the availability of public good and services, ***G***, such as extension services, farmer associations, road quality, etc., because farmers may have access to price information, receive inputs or technical assistance, among other benefits that may affect output.

The vector ***M*** represents farmer's choice of whether to participate or not in the market as a seller, represented by the vector M^{cv}, or as a buyer, represented by the vector M^{cb}. The vector M^{cv} takes value 1 for every crop *c* the farmer decides to sell and $M^{cv} = 0$ for crops not sold. Similarly, the vector M^{cb} takes value 1 for every crop *c* the farmer decides to buy and $M^{cb} = 0$ for crops not bought.[14] Net sales of a particular crop, $NS^{C} \equiv f^{C}(A^{C}, G) - s^{C}$, are positive if and only if $M^{cv} = 1$ and negative if and only if $M^{cb} = 1$. Due to data limitations, the focus of this paper is restricted to comparing farmers' choice as to whether or not to participate in the potato market as a seller.

[14] As mentioned by Boughton *et al.* (2007) and Barrett (2008), households will not both buy and sell the same crop in this one-period model because of the price wedge created by transaction costs. Therefore, there exists a complementary slackness condition, $M^{CV} * M^{CB} = 0$, at any optimum.

The parametric market price each household faces, p^{cm}, is affected by crop-and-household-specific transaction costs, τ^{C} (**A**, **G**, Y, **Z**, $\mathbf{NS}^{\mathbf{C}}$). That is, the household faces wide price margins (i.e. a price band) between the low price at which it could sell a crop and the high price at which it could buy that crop (Sadoulet and de Janvry 1995).[15] These transaction costs create a kinked price schedule, which leads to some households to self-select out of the market for some crops (de Janvry *et al.* 1991; Sadoulet and de Janvry 1995; Boughton *et al.* 2007; Barrett 2008). Following Boughton *et al*. (2007) and Barrett (2008), transaction costs are assumed to be a function of household's productive assets, ***A***, access to public good and services, ***G*** (e.g. good roads and/or participating in farmer organizations may reduce transaction costs), liquidity from off-farm income, *Y*, household-specific characteristics, ***Z***, and amount traded, ***NS***.

The household's choice can be represented by the following optimization problem:

$$\underset{s^c, x,\ A^c, M^{ci}}{Max} \quad U(s^c, x)$$

Subject to the liquidity constraint

$$Y - p^x x + \sum_{c=1}^{C}\left[p^{c*}\left(M^{cv} + M^{cb}\right)\left(f^c\left(A^c, G\right) - s^c\right)\right] = 0$$

And equilibrium conditions for non-tradables

$$A = \sum_{c=1}^{C} A^c$$

$$f^c\left(A^c, G\right) \geq s^c\left(1 - M^{cb}\right) \qquad \text{for c = 1, 2, 3,…, C}$$

With each household-specific crop price determined by the household's net market position:

[15] As mentioned above, shallow local markets and price risks and risk aversion also affect the magnitude of the price band (Sadoulet and de Janvry 1995).

$$p^{c*} = p^{cm} + \tau^{c}\left(A, G, Y, Z, NS^{c}\right) \qquad \text{if } M^{CB} = 1 \text{ (i.e. net buyer)}$$

$$p^{c*} = p^{cm} - \tau^{c}\left(A, G, Y, Z, NS^{c}\right) \qquad \text{if } M^{CV} = 1 \text{ (i.e. net seller)}$$

$$p^{c*} = p^{a} \qquad \text{if } M^{CB} = M^{CV} = 0 \text{ (i.e. autarkic)}$$

Where p^{a} is the autarkic (i.e. non-tradable) shadow price that equates household supply and demand. The second equilibrium condition for non-tradables implies that, if the household does not purchase crop *c* (i.e. $M^{cb} = 0$), production must be greater than or equal to the quantity of crop *c* consumed (may be a net seller) and, if the household does purchase crop *c* (i.e. $M^{cb} = 1$), production must be greater than or equal to zero (may produce crop *c* or not; regardless of which, the household is a net buyer).

To find the optimal solution, two steps are necessary. First, the system must be solved for the optimal solution, conditional on the participation regime (i.e. net seller, net buyer, or autarkic). Then, the market participation regime that yields the highest utility level is chosen (Key *et al*. 2000). That is, the optimal choices of $\{\mathbf{s}^{\mathbf{C}}, \mathbf{A}^{\mathbf{C}}, \mathbf{x}\}$ must be replaced into the utility function to obtain the indirect utility function, ***V***. This indirect utility function must be evaluated under each feasible combination of $\mathbf{M}^{\mathbf{cv}}$ and $\mathbf{M}^{\mathbf{cb}}$ to identify the market participation vectors $\{\mathbf{M}^{\mathbf{cv}}, \mathbf{M}^{\mathbf{cb}}\}$ that yield the highest level of ***V*** (Key *et al*. 2000; Barrett 2008).

Based on the structural model above, the reduced form of each choice variable can be represented as a function of observable (exogenous) variables *A, G, Y, Z,* p^{cm}, and p^{x}. This structural model assumes non-separability[16] in household's production and consumption decisions because the parametric prices are endogenous (because of transaction costs). Because of this, production and consumption behaviors are estimated simultaneously (Sadoulet and de Janvry 1995) in this maximization problem.

Smallholder farmers in rural Angola generally sell their surpluses to itinerant traders at low prices (World Vision 2008). Although this suggests that there may be low barriers to participate in the market; high transaction costs (e.g. obtaining price information, etc.) could make per unit returns to farmers small. Therefore, understanding what factors are affecting smallholder market participation decisions will be useful in designing policies regarding public and private investments oriented to boost market participation by smallholder farmers in rural Angola.

4.2 Econometric Estimation

As mentioned above, this study attempts to estimate the factors associated with household's marketing decisions, focusing on households who sell potatoes in rural Angola. Given that sales are only observed for a subset of the sampled population because farmers who did not sell this crop reported zero sales, the function estimated (i.e. quantity traded) on the selected sample may not estimate the population (i.e. random sample) function (Heckman 1979) due to self-selection problems.[17] Therefore, if the parameters were estimated by least squares, they would be biased and inconsistent (Wooldridge 2009).

---

[16] This implies that production decisions are made as if the household was maximizing profits, while consumption decisions are made as if the household was maximizing utility. For further reading see de Janvry *et al.* (1991) and Sadoulet and de Janvry (1995).

[17] Self-selection arises due to transaction costs, which are reflected in the endogenous market prices faced by farmers.

There are at least three alternatives to least squares to estimate unbiased, consistent and efficient parameters. The *first* alternative is to estimate the parameters using the standard Heckman sample selection model (two step version[18]) used by Goetz (1992), Benfica *et al*. (2006), and Boughton *et al*. (2007). With Heckman two-step approach, one first estimates a probit model of market participation; then, in the second step, one fits a regression of quantity traded by ordinary least squares (OLS), conditional on market participation (Wooldridge 2003). From the probit, one could derive the inverse mills ratio (IMR) and include it as a regressor into the second equation to control for selection bias and obtain unbiased, consistent, and efficient estimators using OLS (for details, see Wooldridge 2003, p. 560-562).

It may seem reasonable that a Heckman selection approach may be appropriate in this context because many households reported zero sales. However, the Heckman regression is designed for incidental truncation, where the zeros are unobserved values (e.g. as with wage rate models where the sample includes unemployed persons) (Ricker-Gilbert *et al.* 2011). Therefore, a corner solution model is more appropriate in this context because, due to market and agronomic conditions, the zeros in the data reflect farmers' optimal choice rather than a missing value (as with Heckman).

The *second* and *third* alternatives to least squares (both corner solution models) are the Tobit estimator proposed by Tobin (1958) and the double hurdle (DH) proposed by Cragg (1971),[19] respectively. Although the Tobit model could be used to model farmers' marketing decisions, its major drawback is that it requires that the decision to sell a particular crop and the decision about how much of that crop to sell be determined by the same process (i.e. the same

[18] Heckman could also be solved by full maximum likelihood (StataCorp 2009).
[19] He proposed a double-hurdle model that nests the usual Tobit model.

variables), which makes it fairly restrictive (Wooldridge 2003 and Ricker-Gilbert *et al.* 2011).[20] In addition to this, in a Tobit model, the partial effects of a particular variable, x_j, on the probability that the farmer will sell and in the expected value of the quantity traded, conditional on participation, have the same signs (Wooldridge 2008).

The DH model is a more flexible alternative (than the Tobit) because it allows for the possibility that factors influencing the decision to sell a crop be different than factors affecting the decision of how much to sell. Therefore, the DH model proposed by Cragg (1971) was implemented in this paper.

In the DH model,[21] the first hurdle estimates the decision of whether or not to participate in the market (i.e. to sell a crop) and, conditional on market participation, the second hurdle estimates the quantity traded (i.e. quantity sold). Due to space limitation, the econometric theory behind the double hurdle model is omitted but it could be made available if requested. In the double hurdle, the decision of whether to sell a crop (a binary variable) is used to estimate the maximum likelihood estimator (MLE) of the first hurdle, which is assumed to follow a probit model. In the second hurdle, the continuous variable of quantity traded is assumed to follow a truncated normal distribution. Therefore, the MLE is obtained by fitting a truncated normal regression model[22] to the quantity traded (Cragg 1971 and Burke 2009). As previously explained, the probability of market participation and the analysis of quantity traded, conditional on market participation, could be determined by different factors (Burke 2009).

---

[20] For details about the Tobit model, see Wooldridge (2003), pg. 540-546.

[21] Also called two-tiered model.

[22] The model is called truncated because the distribution of *y* is truncated at zero to guarantee non-negativity (Cragg 1971).

From the DH model, one could estimate the "unconditional" (on market participation) partial effect (PE) of a particular variable for each observation. Using these PE, one could estimate the average partial effect (APE) of the variable of interest by averaging the PE across all observations in the dataset. However, the standard deviation reported with the ("unconditional") APE should not be used as a standard error for inference about the population because it describes only the data (Burke 2009) and uses an unobservable variable (the IMR from the first hurdle) in its estimation. Instead, two alternatives could be used: (a) standard deviations could be re-estimated by bootstrapping or (b) standard errors could be approximated by the delta method (Burke 2009). In this paper, standard deviations were re-estimated by bootstrapping at 500 repetitions to be able to make inferences about the "unconditional" APEs.

Key *et al.* (2000), showed that, while market participation (i.e. household's decision of whether to sell) depends on both fixed and proportional transactions costs, the quantity supplied, conditional on participation, is only affected by proportional transactions costs. The DH model described above allows for different factors to affect the first and second hurdles, which easily allows excluding fixed costs from the second hurdle. However, the variables used as proxies for fixed costs (i.e. distance to market and quality of the road) were included in both the market-participation and the quantity-traded regressions to test whether fixed costs only affect the first hurdle among Angolan farmers.

Although the independent variables included in the regressions are explained in the next section, the quantity harvested (included in both hurdles) is worth discussing here. Quantity harvested is potentially endogenous to the decision of whether to participate in the market as a seller and on the decision of how much to sell. For instance, if a farmer produces a crop with the intention of selling his surplus, whether he participates in the market will depend on how much he harvests--i.e. if the quantity harvested is small, he might decide to keep his production for his own consumption. Furthermore, market conditions will influence the amount a farmer produces because if the farmer perceives that he could sell in the market, he may decide to produce more for this purpose. Because of all these, there may be correlation between the error term in a reduced equation of quantity harvested and the error term of the probability of participation and quantity traded; thus, making quantity produced an endogenous covariate.

To deal with this potentially endogenous variable, an OLS regression was estimated on the quantity produced. Then, the residuals from this OLS regression were estimated and included in both the probit and truncated normal regressions as an additional explanatory variable. This allowed to test whether quantity produced was truly endogenous (i.e. if the coefficient of this variable is statistically significant, quantity produced is endogenous). Although several variables included in the OLS regression were also in the DH estimation, the former model included additional variables that were not expected to affect marketing decisions.

5 Data Used

Data used in this study came from the cross sectional household- and village-level survey implemented by World Vision's ProRenda project in Angola in 2009. World Vision, in collaboration with ACDI/VOCA,[23] the Ministry of Agriculture and Rural Development of Angola, the Angolan NGO HORIZONTE, and Michigan State University are implementing a four-year project[24] targeted at increasing smallholder-farming families' annual income from non-perishable crops (World Vision 2008). The ProRenda project attempts to increase

[23] Agricultural Cooperative Development International / Volunteers in Overseas Cooperative Assistance.

[24] The ProRenda Project, which is financed by the Bill and Melinda Gates Foundation.

smallholder's (60% of the beneficiaries will be women) incomes by establishing competitive value chains for potatoes, beans, onions and other high-value crops.

The baseline survey was implemented from January through April of 2009 and collected data about the latest harvest between September 2007 and December 2008. In Angola, the agricultural year goes from September through May of the following year (MINADER and FAO 2003). Therefore, the data collected refers to the 2007-2008 agricultural year and the first season of the 2008-2009 agricultural year.

The survey was implemented in three provinces of the central highlands of Angola: Huambo, Bie, and Bengela. These provinces have the most productive lands within the highland region (World Vision 2008) because of good rainfall distribution and environmental conditions; however, yields are usually low (MINADER and FAO 2003). The major crops produced in the highlands are: corn, wheat, rice, potatoes, sweet potatoes, beans, cassava, sugarcane, peanuts, sunflower, sesame, tobacco, and vegetables (MINADER and FAO 2003).

The survey included a total of 656 households[25] across 40 communities. The households were selected using a clustering sampling methodology. This means that the villages were selected first; then, within those villages, the households were selected. While the villages were selected systematically using probability proportional to size, the households were classified into four categories (based on gender of household head and participation in farmer organization) and, within each category, a random systematic sample of households was selected.[26] In order for the sample estimates to be representative of the population covered by the survey, sampling

[25] However, only 620 surveys were valid and used in the analysis.

[26] Details about the sampling methodology and weight estimation can be found in Reyes *et al*. (2010).

weights were used. The basic weight for each sampled household is the inverse of its probability of selection (see Reyes *et al*. 2010 for details).

The household-level survey collected information about households' socioeconomic characteristics, productive and non-productive assets, participation in farmer organizations, and production and marketing information of beans, potatoes, onions, carrots and cabbages. The village-level survey collected information regarding the distance between the village and the main commercial town (or "sede"), availability of public services (e.g. telephone, electricity, banks, health clinics, local markets) and public transportation, and quality of the road between the village and the main commercial town.

The independent variables included in the regressions were classified into five categories: (1) household characteristics, (2) private assets, (3) public assets and quasi-fixed factors, (4) production- and marketing-related variables, and (5) squared and interaction terms (Table 5.1).[27] These variables were included because were theoretically expected to affect production and marketing decisions. A total of 40 independent variables were used to estimate the three models proposed in the previous section: linear regression model of quantity produced, probit model of market participation, and truncated normal regression model of quantity traded.

[27] This last category was only used in the OLS regression of quantity produced.

Table 5.1. Independent variables included in the production and marketing decision regressions. Angola, 2009.

No.	Variable	Model where included[1]
	Dependent:	
	Quantity produced (kg)	1
	Market participation (1=yes)	2
	Quantity sold (kg)	3
	Household (HH) Characteristics:	
1	Age of HH head (yr)	1, 2, 3
2	Gender of HH head (1=male)	1, 2, 3
3	Dependency ratio	1, 2, 3
4	HH member is in farmer organization (1=yes)	1, 2, 3
5	No. adults who can read & write	1, 2, 3
6	No. of tropical livestock units (TLU) owned	1, 2, 3
	Private Assets (1=yes):	
7	Own plow	1
8	Own backpack sprayer	1
9	Own motorcycle	2, 3
10	Own bicycle	2, 3
11	Index of home and transportation assets a/	1
12	Index of home assets b/	2, 3
13	Index of productive assets c/	2, 3
	Public Assets and Quasi-fixed Factors:	
14	IDA office in the village (1=yes)	1, 2, 3
15	Public market available in the village (1=yes)	1, 2, 3
16-22	Seven dummy variables for municipalities (1=yes)	1, 2, 3
23	Distance from village to commercial town (km)	2, 3
24	Road between village and commercial town in poor condition (1=yes)	2, 3
	Production- and Marketing-related Variables:	
25	Seed used (kg)	1
26	Type of plot (1=rainfed plot)	1
27	Planted seed of local variety (1=yes)	1
28	Used fertilizer (1=yes)	1
29	Used pesticides (1=yes)	1
30	Reported production costs (Kw/kg)	1
31	HH reports lower harvest (1=yes)	1
32	Seller sought price information prior to sales (1=yes)	3
33	Reported marketing costs (Kw/kg)	3
34	Quantity produced (kg)	2, 3

Table 5.1 (cont'd).

No.	Variable	Model where included[1]
	Squared and interaction terms:	
35	Age of HH head squared	1
36	Dependency ratio squared	1
37	No. adults who read & write squared	1
38	TLU squared	1
39	Seed used squared	1
40	Production costs * HH reported lower harvest	1

[1] Model 1 = Ordinary Least Squares for production; Model 2 = Probit for market participation; Model 3 = Truncated Normal Regression for quantity sold.
NOTES:
a/ Index of home and transportation assets include ownership of cell phone, television, radio, having a latrine in the homestead, having a roof made of improved materials (e.g., zinc), having a water storage facility at home, ownership of motorcycle, and ownership of bicycle.
b/ Index of home assets include the same assets mentioned in "a/" excluding owning a motorcycle and/or a bicycle.
c/ Index of productive assets include ownership of plow, cart, and backpack sprayer.

Although most variables are self-explanatory, a brief explanation of key variables is provided next. The dependency ratio was estimated by dividing the number of people younger than or equal to 17 by the household size. Having a household member participating in farmer organizations refers to any member of the household who participated in FO within the previous 12 months. Adult literacy refers to members older than 17 who can read and write. The number of tropical livestock units was estimated using FAO conversion factors for South Africa where, for example, one cattle equals 0.70 livestock units; one sheep equals 0.10 livestock units, etc. It included oxen, cattle, goats, sheep, pigs, chicken, and rabbits.

An asset index was estimated to classify households according to its (asset) wealth and was used as a proxy for household wealth. Details are included in section 6.2. The quasi-fixed variables included having an IDA (the government's Institute for Agrarian Development) office in the village, access to public markets for consumption, and seven dummies[28] for the municipalities where the households were located to control for variations in environment and marketing conditions faced by farmers (at the macro-level).[29] Fixed transaction costs (FC) included the distance between the village and the main commercial town (or "sede") and the quality of the road between these two places.

The production-related variables are self-explanatory except for one--type of plot. Angolan farmers in these provinces could cultivate in one (or several) of four possible types of plots: *nacas*, *ombandas*, *otchumbo, and lavras*.[30] *Nacas* are irrigated lowland areas located close to river deltas, used during the dry season (by exploiting residual moisture), and account for 4% of the cultivated area. *Ombandas* are medium-level lands with access to gravity-fed irrigation, used in all seasons, and account for 15% of the cultivated area. *Otchumbo* are small areas close to the homestead, intensively cultivated all year round, and account for 4% of the cultivated area. Finally, *lavras* are upland areas used for rainfed agriculture and account for 77% of the cultivated area (World Vision 2008). Given that *lavras* are the most commonly used types of plots, a dummy variable was created to account for whether the crop was produced in this type of plot.

Unit production costs were obtained by adding reported costs on fertilizers, seed, pesticides, labor, and transport from the field to the home and dividing this by total quantity produced. Similarly, unit marketing costs were obtained by adding farmers' reported costs of use of bags, sewing of these bags, transportation costs, loading and unloading of the output, and

---

[28] The dummy for Londuimbali municipality was excluded to avoid the dummy variable trap.
[29] Although it would have been ideal to include dummy variables for each community, this was not practical because there were 40 communities.
[30] These are Portuguese names with no English translation.

taxes and fees paid at the market and dividing this by total quantity sold. The squared terms were included to allow for non-linear relationships between independent and dependent variables only in the OLS regression. Finally, the residuals of the OLS regression (on quantity produced) were included in both hurdles to test for endogeneity of this variable.

6 Results

This section is divided into three subsections. The first subsection describes the sample and provides the socioeconomic characteristics of farm families, focusing on the variables of interest for the double hurdle analysis and the results are disaggregated by market participation. The second subsection briefly describes the OLS regression results of the quantity produced. The last subsection details the double hurdle regression results. Before discussing the results, it is worth explaining how the asset indexes were estimated.

Although the details are not presented in this paper due to space limitation, several asset indexes were estimated using primary component analysis. First, a general asset index was estimated considering ownership of tractors, trucks, cars, plows, carts, backpack sprayers, motorcycles, bicycles, cell phones, radio, televisions, water storage facilities and latrine at the homestead, and whether the roof was made of zinc or lusalite (considered improved materials). However, tractors, trucks and cars were excluded because no household in the sample owned these items. The assets with the highest "weight" (i.e. more importance) in the index were: owning a television, a cart, a motorcycle and a cell phone. In contrast, the asset with the lowest "weight" in the index was having a latrine at home (since most farmers had a latrine at home). This index (used as a proxy for wealth) suggests that male-headed households were richer than female-headed ones (the average index for male heads was 0.512 vs. -0.625 for female heads), farmers growing potatoes were richer than non-growers (index = 0.269 vs. -0.037 for non-growers), and that, within potato producers, sellers were richer than non-sellers (index = 0.432 vs. -0.118 for non-sellers). These results are confirmed by a graphical analysis of the cumulative distribution of the index by gender of the head (Figure 6.1) and crop grown (Figure Annex 6.1).

Second, three additional indices were estimated using the same eleven assets included in the general index: (1) an index of home assets, which included all assets except ownership of plows, carts, backpack sprayers, motorcycles, and bicycles; (2) an index of productive assets, which only included ownership of plows, carts, and backpack sprayers, and excluded all other assets; and (3) an index of home and transportation assets (which included all assets except ownership of plows, carts, and backpack sprayers). This allowed evaluating the effect of productive assets separately from other non-productive and home assets.

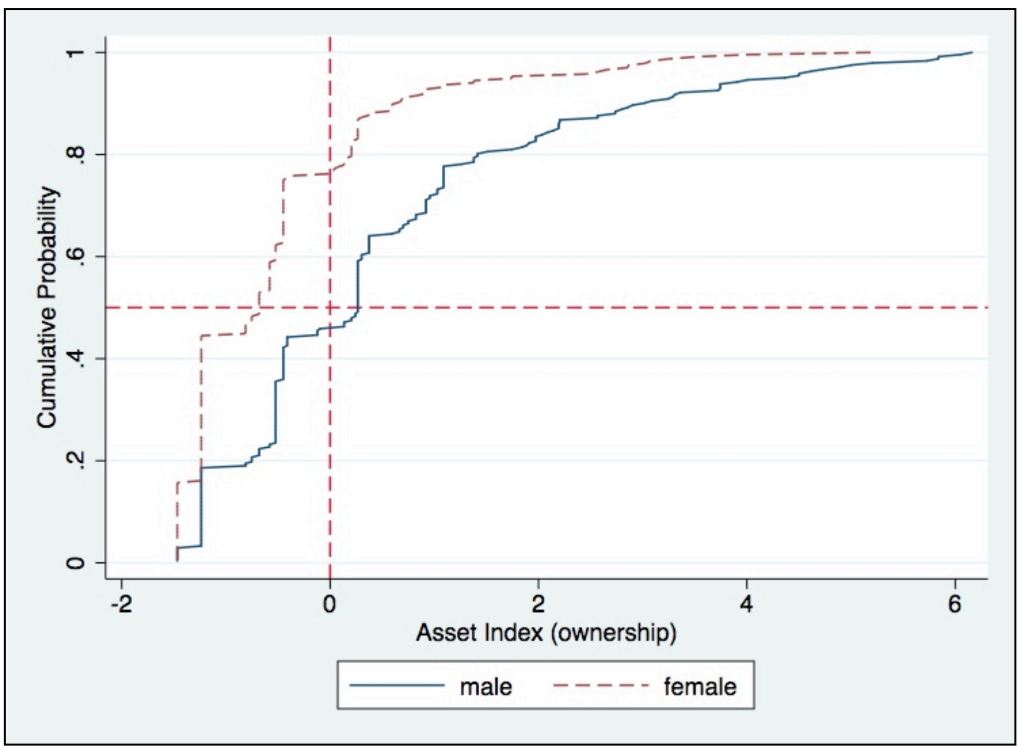

Figure 6.1. Cumulative distribution of asset index by gender of household head. Central Highlands of Angola, 2009.

6.1 Descriptive Statistics

Potatoes were planted by 55% of the farmers in the region.[31] Cultivating potatoes was common among richer farmers (as classified by the general asset index;[32] Figure Annex 6.1) and male-headed households. On average, each farmer sold 200 kg of potatoes, which corresponds to roughly 87% of sellers' production (Table 6.1). Furthermore, farmers who didn't sell produced less than sellers: non-sellers produced only 13% of what potato sellers did (Table 6.1).

The differences in age of the head between sellers and non-sellers were not statistically significant at the 10% level (Table 6.1). As expected, more male-headed households participated in markets as sellers (1% significance level, SL). Furthermore, there were slightly more than one dependent for every two adults in the household (the average dependency ratio was 0.56) and potato sellers had significantly more dependents than non-sellers. The share of households having at least one member participating in farmer organizations (FO) in the year prior to the interview was significantly higher for potato sellers (Table 6.1).

Not surprisingly, the index of home assets suggests that potato sellers owned more home assets than non-sellers (1% SL). In contrast, although the index of productive assets was higher for non-sellers, the differences were not statistically significant at the 10% level. Furthermore, the differences on access to public assets (i.e. IDA office and public market) were not statistically significant between potato sellers and non-sellers (Table 6.1).

[31] Furthermore, 71% of farmers planted beans and 46% planted onions. Although these crops are not included in this study, the data collected included information about these crops.

[32] Asset index and economic status index are used interchangeable. This index was estimated using primary component analysis, following Filmer and Pritchett (2001), McKenzie (2005) and Reyes *et al*. (2010).

Table 6.1. Descriptive statistics of the variables used in the Double Hurdle analysis. Central Highlands of Angola, 2009.

	Potato					
	Non-sellers		Sellers			
Demographics	Mean	S.E.	Mean	S.E.	MT[1]	**Total**
Quantity sold (kg)	n.a.		200	22.36	--	
Household Characteristics						
Age of head (years)	42	3.858	39	0.411		40
Gender of head (% male)	52	0.283	78	0.195	***	71
Dependency ratio[2]	0.50	0.018	0.58	0.022	**	0.56
HH member is in FO[3] (% yes)	4	0.025	11	0.054	*	10
Family members >17 who are literate[4]	0.9	0.198	0.7	0.085		0.8
No. of Tropical Livestock Units[5]	0.47	0.121	0.36	0.083		0.39
Owns motorcycle (% yes)	4	0.022	10	0.026		9
Owns bicycle (% yes)	25	0.155	29	0.061		28
Index of home assets	-0.27	0.130	0.35	0.211	***	0.19
Index of productive assets	0.20	0.336	0.07	0.061		0.11
Public Assets and Quasi-fixed factors						
IDA[6] office in village (% yes)	17	0.052	26	0.040		23
Public market available in village (% yes)	19	0.052	16	0.052		17
Mean[7] sales price, local market (kw/kg)	88.4	5.246	75.1	2.995	***	78.6
Percent of HH in following municipalities:						
Caala	23	0.065	11	0.020	**	14
Ekunha	1	0.008	2	0.017		2
Bailundo	21	0.062	19	0.053		19
Londuimbali	35	0.054	15	0.030	***	21
Katchiungo	4	0.014	15	0.021	**	12
Tchicalachuluanga	7	0.029	2	0.016	**	3
Chiguar	9	0.041	36	0.059	***	29
Babaera	0.6	0.005	0.2	0.002		0.3
Distance from village to sede (km)	10.3	0.990	11.4	0.908		11.1
Road between village and sede in poor condition[8] (% yes)	66	0.100	81	0.029	**	77
Production and Marketing variables						
Quantity produced (kg):	30	8.85	230	29.72	***	177
In Caala	50	16.02	359	59.52	--	225
In Ekunha	66	0.00	417	54.89	--	374
In Bailundo	12	1.91	42	11.47	--	33
In Londuimbali	30	15.44	172	26.64	--	108
In Katchiungo	20	11.34	224	51.70	--	206

Table 6.1 (cont'd).

Demographics	Potato Non-sellers Mean	S.E.	Potato Sellers Mean	S.E.	MT[1]	Total
In Tchicalachuluanga	16	2.66	136	15.89	--	71
In Chiguar	37	11.36	312	61.55	--	291
In Babaera	12	3.12	92	9.64	--	50
Seller sought price information prior to sales (% yes)	n.a.		63	0.045	--	
Reported marketing costs (Kw/kg)	n.a.		2.9	0.249	--	
Number of observations	75		165			240

[1] MT = test of difference between means: *significant at 10%; **significant at 5%; ***significant at 1%; -- not tested; n.a. = not applicable.

[2] Dependency ratio estimated by dividing the number of people 17 years or younger by the household size.

[3] FO = Farmer organization.

[4] Literacy refers to people who can read and write.

[5] Tropical Livestock Units estimated using FAO conversion factors.

[6] IDA = Government's Institute for Agrarian Development.

[7] For farmers who sold in local markets, their reported price was averaged per community. Communities with missing prices use average price per the next political division (i.e. town, municipality).

[8] Poor condition means the road is a dirt road, not rehabilitated (i.e. without maintenance).

Source: ProRenda survey, Angola, 2009. Estimates weighted to reflect population.

As previously explained, sales prices were collected for farmers who sold (part of) their output. Farmers reported selling their output in different places, including their farm, their home, local markets and other markets. To control for (potential) endogeneity problems in market prices, for farmers who reported selling at local markets, the average sale price was estimated. However, in some villages, none of the farmers who sold their output did so in local markets; thus, the average price could not be estimated. In these cases, the average price of the next political division (i.e. town, municipality) was estimated. Although this information is presented in Table 6.1, it was excluded from the double hurdle analysis because it was judged to be inaccurate. That is, since prices were imputed to non-sellers, in some villages, a high share of non-sellers were imputed a high price, thus offsetting any positive effect of this variable.[33]

The highest share of potato sellers was in the Chiguar municipality. However, the highest production was distributed among Caala, Ekunha and Chinguar municipalities. The average distance between the villages and their main commercial town was 11.1 km (Table 6.1). In general, a higher percent of sellers were located in villages farther away than non-sellers; thus, the average distance from their villages to their main commercial town was higher for sellers. For example, while 35% of potato non-sellers were located in Londuimbali (with an average of 6.6 km), a similar percent of potato sellers were located in Chinguar where, villages were located farther away (14.3 km) from their main commercial town. Furthermore, a higher share of potato sellers was located in villages with poor road conditions between the village and the main commercial town (Table 6.1). Finally, less than two-thirds of the farmers who sold potatoes obtained price information before selling their surpluses and sellers reported an average marketing cost of 2.9 Kwanzas[34] per kilogram sold (Table 6.1).

6.2 Econometric estimation of factors influencing potato production

It was suspected that production could be an endogenous covariate in the double hurdle analysis. Thus, a linear regression (OLS) estimation was used to determine which factors were affecting potato production. Then, the residuals of this regression were included as an additional

[33] When included in the double hurdle regressions, this variable was statistically not significant or had a negative sign, which is contrary to what economic theory suggests.
[34] The exchange rate at the time of the survey was 75 Angolan Kwanzas per US$.

explanatory variable in the double hurdle analysis and tested for endogeneity. This subsection presents the results of the OLS regression on quantity produced.

The descriptive results of the factors influencing production are included in Table Annex 6.1. Interested readers can refer to this table for details. The econometric results of the OLS regression are presented in Table 6.2. The model appears to slightly over fit the data since its R-squared is 0.7. The results show that male heads produced, on average, 55 kg more than female heads (1% SL). Thus, providing technical assistance (related to production) to female-headed households may be necessary to help them obtain higher production. Surprisingly, none of the productive assets or access to public goods (i.e., IDA office in the village or public market available in the village) had a statistically significant effect on production.

Farmers producing potatoes in Caala, Ekunha, Tchicalachuluanga, and Chinguar municipalities produced statistically more potatoes than farmers in the Londuimbali municipality. The differences in production between all other municipalities and Londuimbali were not statistically significant at the 10% level.

Most production-related variables had statistically significant effects on production (Table 6.2). Since the dependent variable in this model was production (not yields), it was expected that, as seed use increased, quantity produced would increase. Thus, the finding that quantity produced was positively affected by the amount of seed used was no surprise. Although farmers using local varieties obtained lower production, the differences between farmers who used local varieties and farmers who used improved varieties were not statistically significant at the 10% level (Table 6.2).

Table 6.2. Linear regression model of factors influencing potato production (kg). Central Highlands of Angola, 2009.

	N = 264	
	R-squared = 0.7000	
Independent variables	Coefficient	p-value
Household (HH) Characteristics		
Age of HH head (Years)	-3.84	0.109
Gender of HH head (1=Male)	54.59	***0.003
Dependency ratio[1]	-384.72	0.153
HH member is in farmer organization (1=Yes)	-10.05	0.766
No. adults (>17 yr) literate[2]	18.28	0.248
No. of Tropical Livestock Units	-9.01	0.758
Index of home and transportation assets	16.39	0.217
Productive Assets Ownership (1=Yes)		
Owns a plow	43.08	0.521
Owns a backpack sprayer	-12.42	0.612
Public Assets and Quasi-fixed Factors (1=Yes)		
IDA office in village	7.24	0.577
Public market in village	36.64	0.175
HH in Caala Municipality	111.66	**0.049
HH in Ekunha Municipality	220.40	**0.015
HH in Bailundo Municipality	32.31	0.466
HH in Katchiungo Municipality	109.38	0.110
HH in Tchicalachuluanga Municipality	91.50	*0.092
HH in Chiguar Municipality	116.59	***0.002
HH in Babaera Municipality	91.38	0.138
Production-related variables		
Total seed used (kg)	1.65	**0.018
Planted in rainfed plot (1=Yes)	-13.47	0.426
Planted local variety (1=Yes)	-41.23	0.168
Used fertilizer (1=Yes)	44.74	***0.004
Used pesticides (1=Yes)	74.87	0.130
Reported production costs (Kw/kg)	-0.92	**0.020
HH reported lower harvest (1=Yes)	-67.73	**0.040
Squared and interaction terms		
Age squared	0.02	0.370
Dependency ratio squared	555.36	0.102
No. adults literate squared	-5.60	0.423
Tropical Livestock Units squared	4.80	0.617
Total seed used squared	0.004	*0.064
Production costs * HH reported lower harvest	0.72	0.109

Table 6.2 (cont'd).

Independent variables	N = 264 R-squared = 0.7000 Coefficient	p-value
Constant	126.89	**0.028

Notes: *, **, *** indicates the corresponding coefficients are significant at the 10%, 5%, and 1% levels, respectively. All municipalities compared to Londuimbali municipality.

[1] Dependency ratio estimated by dividing No. members ≤17 yr by household size.

[2] Literacy refers to adults who can read and write.

Source: ProRenda survey, Angola, 2009. Estimates weighted to reflect population.

As expected, use of fertilizer positively affected production. Farmers who used fertilizer obtained, on average, 45 kg more potatoes than farmers who didn't apply fertilizer to their fields (Table 6.2). Not surprisingly, as the per unit production cost increased, quantity produced decreased. Although these results do not suggest whether farmers didn't have access to fertilizer or could not afford to buy this input (i.e. due to high price), given that the largest share of production costs were due to expenses in fertilizer and that 65% of farmers applied fertilizer, is likely that most farmers could not afford to purchase the required amounts of fertilizer.

Finally, farmers reporting that their current harvest was lower than their harvest in a normal year obtained lower production. Approximately 60% of farmers who reported that their current harvest was lower mentioned the little or no use of fertilizer as the reason for this.

6.3 Double hurdle regression results

The descriptive statistics for the variables included in these models were already discussed at the beginning of this section; therefore, this subsection dives directly to the double hurdle regression results. One additional point is worth discussing here. The coefficient of the residuals of the OLS regression on production was not statistically significant in any of the two hurdle regressions (p-value = 0.785 for hurdle 1 and p-value=0.357 for hurdle 2). These results suggest that quantity produced was not endogenous; therefore, the residuals were excluded from both hurdles and quantity produced was treated as an exogenous covariate.

The double hurdle regression results are presented in Table 6.3. While male-headed households were more likely to participate in the market as sellers (5% SL), once the market participation decision has been made, gender of the head had no significant effect on the quantity of potatoes sold. Targeting assistance to female heads may be necessary to increase their participation in the potato market as sellers, which would benefit them due to increased income from potato sales.

Having more dependents increases the likelihood of market participation (10% SL). In contrast, as the number of literate adults in the household increase, farmers are less likely to sell potatoes (5% SL). However, after the participation decision has been made, having more dependents or more literate adults in the household have no statistically significant effect on the ("conditional") amount of potatoes sold (Table 6.3).

While owning a bicycle was not associated with the likelihood of selling potatoes, owning this low-cost transportation vehicle (conditional on market participation) was positively associated with the quantity of potatoes sold (5% SL). This was perhaps due to the fact that a bicycle could easily be used to transport potatoes to local markets or other places for sale.

While the probability of selling potatoes was positively associated with the index of home assets (1% SL), conditional on selling potatoes, the quantity sold was negatively associated with this index (1% SL; Table 6.3). These results suggest that, although richer households (per the home assets they own) were more likely to participate in the market as sellers, they sold fewer potatoes.

Table 6.3. Double-Hurdle model of factors influencing potato marketing decisions. Central Highlands of Angola, 2009.

	HURDLE 1		HURDLE 2	
	Probability of selling Probit Estimator		Quantity sold (kg) Truncated Normal Regression Estimator	
	N = 240		N = 159	
Independent variables: the coefficients displayed are the conditional average partial effects (APEs).	Pseudo R2 = 0.5085		Prob > Chi2 = 0.000	
	Coefficient	p-value	Coefficient	p-value
Age of HH head (Years)	-0.0007	0.700	-0.420	0.387
Gender of HH head (1=Male)	0.115	**0.035	21.227	0.136
Dependency ratio	0.188	*0.083	8.237	0.808
HH member is in farmer organization (1=Yes)	0.074	0.117	23.432	0.216
No. adults (>17 yr) literate	-0.069	**0.015	6.815	0.355
No. of Tropical Livestock Units	-0.067	0.116	-8.254	0.385
Owns motorcycle (1=Yes)	-0.073	0.523	18.260	0.667
Owns bicycle (1=Yes)	-0.068	0.270	38.367	**0.019
Index of home assets	0.061	***0.007	-13.987	***0.006
Index of productive assets	0.011	0.434	16.828	**0.043
IDA office in village (1=Yes)	0.098	0.131	47.555	***0.007
Public market in village (1=Yes)	-0.182	**0.018	-14.219	0.344
HH in Caala Municipality (1=Yes)	-0.171	0.138	56.428	0.154
HH in Ekunha Municipality (1=Yes)	-0.045	0.714	64.322	*0.081
HH in Bailundo Municipality (1=Yes)	0.165	**0.032	-101.602	**0.040
HH in Katchiungo Municipality (1=Yes)	0.187	***0.008	12.279	0.594
HH in Tchicalachuluanga Municipality (1=Yes)	-0.054	0.582	-14.295	0.508
HH in Chiguar Municipality (1=Yes)	0.173	**0.034	-14.091	0.587
HH in Babaera Municipality (1=Yes)	-0.079	0.553	-68.719	***0.004
Distance from village to sede (km)	0.003	0.338	-0.929	0.476
Road between village and sede in poor condition (1=Yes)	-0.020	0.726	-64.770	**0.016
Seller sought price information prior to sales (1=Yes)	n.a.		-4.390	0.755
Reported marketing costs (Kw/kg)	n.a.		-0.611	0.517
Total potato production (kg)	0.002	***0.000	0.538	***0.000

Notes: *, **, *** indicates the corresponding coefficients are significant at the 10%, 5%, and 1% levels, respectively. Coefficients and p-values obtained using the *margins* command in Stata.

Dependency ratio estimated by dividing No. members <17 yr by household (HH) size. Literacy refers to adults who can read and write. n.a. = not applicable because variable was not included in the regression.

Source: ProRenda survey, Angola, 2009. Estimates weighted to reflect population.

In contrast, while market participation was not statistically associated with the index of productive assets, the conditional quantity sold was positively affected by owning productive assets (5% SL). This suggests that, conditional on participation, farmers who own productive assets sell more potatoes.

Although having a government's extension office in the village had no statistically significant effect on the likelihood of becoming a seller, having this extension service positively affect the quantity of potatoes sold (1% SL). In contrast, having access to a public market for purchasing food or selling surpluses was a marginally significant negative factor in market participation decision. The main reason for this may be the fact that a higher share of non-sellers reported that public markets were available in their villages (19% vs. 16% sellers.) However, once the market participation decision has been made, this factor had no statistically significant effect on the quantity of potatoes sold.

In contrast to Key *et al*. (2000) the results suggest that one of the proxies for fixed costs had a statistically significant (5% SL) negative effect on the quantity of potatoes sold. Farmers located in villages with poor road quality between the village and the main commercial town sold fewer potatoes. Although only 33% of farmers reported selling at least one of their outputs in other markets (i.e. outside the village, for whom road quality may be important), these farmers sold more than double the amount sold by farmers selling at home or in the local market (308 kg vs. 145 kg, 1% SL.) Thus, investing in improving roads could be an important factor to boost potato sales.

Finally, production was a marginally significant positive factor on both the probability of market participation and quantity traded. This was expected since farmers who have greater production have more surpluses they could sell. Although the magnitude on the market participation decision is small, approximately 54% of increased production, conditional on market participation, would be sold (i.e. for every extra kg produced, 0.54 kg would be sold).

The unconditional (on market participation) average partial effects (APE) of all variables are included in Table Annex 6.2. The APE incorporates the partial effect of both hurdles, which allows making unconditional inferences about the factors affecting the quantity of potatoes sold. Although male-headed households sold more potatoes, the differences between these households and female-headed households were not statistically significant at the 10% level. Thus, the unconditional quantity of potatoes sold was gender neutral. This may be explained by the fact that 41% of female-headed households reported their (male) spouses as the ones responsible for sales (vs. 26% of male-headed households reporting female spouses as responsible for sales.) This result suggests that households led by (married) females rely on their (male) spouses for marketing-related decisions; therefore, explaining why the differences in the quantity sold were not statistically significant.

In contrast to the conditional quantity of potatoes sold, having a member of the household participating in a farmer organization was positively associated (5% SL) with the unconditional quantity of potatoes sold (Table Annex 6.2). Thus, promoting participation in these organizations or establishing farmer organizations in villages without them could boost potato sales.

Owning a bicycle was positively correlated with the unconditional quantity of potatoes sold (1% SL). As previously discussed, a bicycle could easily be used to transport potatoes for sale. Further, owning more home assets (as indicated by the index of home assets) was a statistically significant negative factor in the unconditional quantity of potatoes sold (5% SL). However, the magnitude of this effect was very small--an increase of one unit in this index would reduce the quantity sold by approximately 9 kg. The finding that these potentially richer farmers (who have more home assets) sell fewer potatoes may be explained by the fact that a lower percent (39%) of richer potato producers reported potatoes as the major source of crop income (compared to 43% of farmers in the poorest tercile) and because a larger percent (30%) of richer potato producers reported services as the main source of non-crop income (compared to 1% of farmers in the poorest tercile). Thus, richer farmers have diverse sources of income, which make them less dependent on potato sales.

The presence of an IDA office in the village was positively correlated with the unconditional quantity of potato sold (Table Annex 6.2). Farmers in villages with IDA offices sold, on average, 46 kg more potatoes than farmers in villages without IDA offices. Thus, providing farmers with extension services could contribute to increase potato sales.

Not surprisingly, farmers located in villages with poor road quality between the village and the main commercial town sold fewer potatoes (1% SL). Lastly, if production increases by one kilogram, 57% of this increase will end up being sold (1% SL). Therefore, investing in public infrastructure (i.e., improving roads) and devoting efforts targeted at helping farmers increase their production could positively affect the "unconditional" quantity of potatoes sold.

The results suggest that, to boost the unconditional quantity of potatoes sold by smallholder farmers in the central highlands of Angola, investments are needed to (a) promote farmer participation in organizations and/or establish farmer organizations in villages without them, (b) provide assistance to poorer potato producers (as classified by the index of home assets); however, since this crop requires investments, this assistance can't focus on farmers who are too poor, (c) provide extension services related to both production and marketing aspects to farmers, (d) invest in infrastructure, especially in improving the quality of the roads, and (e) help farmers increase their potato production, which can be done by making inputs more affordable and/or available and by targeting assistance to female-headed household, who produce less than their counterparts.

7 Conclusion

This paper uses a double hurdle regression analysis to estimate the factors influencing marketing decisions among potato growers, focusing on gender of head, asset ownership and transaction costs. Although the results suggest that the quantity produced is exogenous in the models for market participation and for quantity sold, the methodology used provides a framework for others to follow when endogeneity is suspected in one or more variables.

The wealth analysis (using an asset index) suggests that potato growers, potato sellers and male heads are richer than their counterparts. The linear regression results on quantity produced suggest that female-headed households produce less than their male counterparts. Thus, providing technical assistance to female-headed households may be necessary to help them obtain higher production. Furthermore, owning productive assets had no statistical effect on production perhaps due to the fact that only a small share of farmers owned these assets. Moreover, all production-related variables were affecting production in the expected way (although not all were statistically significant). The results suggest that farmers who use fertilizer produce more than farmers who don't apply fertilizer to their fields. Surprisingly, there was no statistically significant effect of using local varieties vs. using improved varieties on production.

The results of the double hurdle analysis suggest that male-headed households are more likely to sell potatoes. However, once the market participation decision has been made, the conditional quantity sold is gender neutral. Furthermore, while owning more home assets positively affect the probability of selling potatoes, once this decision has been made, owning more home assets negatively affect the quantity sold. In contrast, owning productive assets have a significantly positive effect on the conditional quantity of potatoes sold.

Although having a government's extension office (IDA office) in the village had no significant effect on the likelihood of market participation, farmers in villages with IDA offices sold more potatoes than farmers without this public service. In contrast, having a public market for consumption in the village negatively affected market participation. As expected, having a poor quality road between the village and main commercial town negatively affected the conditional quantity of potatoes sold. Finally, as the quantity produced increased, the likelihood of selling potatoes and, conditional on participation, the quantity sold both increased.

The unconditional APEs suggest that, to boost the unconditional quantity of potatoes sold by smallholder farmers in the central highlands of Angola, investments are needed to (a) promote farmer participation in organizations and/or establish farmer organizations in villages without them, (b) provide assistance to poorer potato producers; however, since this crop requires investments, this assistance can't focus on farmers who are too poor, (c) provide extension services related to both production and marketing aspects to farmers, (d) invest in infrastructure, especially in improving the quality of the roads, and (e) help farmers increase their potato production, which can be done by making inputs more affordable and/or available and by targeting assistance to female-headed household, who produce less than their counterparts. Thus, boosting potato sales would be a challenge for the government of Angola and donors since, due to its strong currency, overcoming these limiting factors may require large amounts of financial and human resources.

ANNEXES

Table Annex 6.1. Descriptive statistics of factors influencing potato production. Central Highlands of Angola, 2009.

Variables	Potato (N=264)	
	Mean	Linearized Std. Err.
Dependent Variable: Quantity produced (kg)	153.88	23.702
Independent Variables		
Household (HH) Characteristics		
Age of HH head (Years)	39.16	0.843
Gender of HH head (% Male)	0.72	0.226
Dependency ratio[1]	0.59	0.010
HH member is in farmer organization (% yes)	0.10	0.046
No. adults (>17 yr) literate[2]	0.82	0.069
No. of Tropical Livestock Units	0.35	0.070
Index of home and transportation assets	0.32	0.260
Productive Assets Ownership (% yes)		
Owns a plow	0.12	0.025
Owns a backpack sprayer	0.04	0.014
Public Assets and Quasi-fixed Factors (% yes)		
IDA office in village	0.27	0.045
Public market in village	0.17	0.043
HH in Caala Municipality	0.14	0.017
HH in Ekunha Municipality	0.02	0.012
HH in Bailundo Municipality	0.16	0.040
HH in Londuimbali Municipality	0.18	0.041
HH in Katchiungo Municipality	0.12	0.023
HH in Tchicalachuluanga Municipality	0.03	0.019
HH in Chiguar Municipality	0.34	0.047
HH in Babaera Municipality	0.00	0.002
Production-related variables		
Total seed used (kg)	35.45	4.380
Planted in rainfed plot (% yes)	0.43	0.018
Planted local variety (% yes)	0.75	0.054
Used fertilizer (% yes)	0.65	0.030
Used pesticides (% yes)	0.10	0.050
Reported production costs (Kw/kg)	63.11	11.118
HH reported lower harvest (% yes)	0.66	0.022

[1] Dependency ratio estimated by dividing No. members ≤17 yr by household size.

[2] Literacy refers to adults who can read and write.

Source: ProRenda survey, Angola, 2009. Estimates weighted to reflect population.

Table Annex 6.2. Unconditional average partial effects of factors influencing potato sales. Central Highlands of Angola, 2009.

Independent variables: the coefficients displayed are the unconditional average partial effects (APEs).	**Quantity sold (kg)**	
	Coefficient	p-value
Age of HH head (Years)	-0.392	0.133
Gender of HH head (1=Male)	23.189	0.355
Dependency ratio	15.323	0.790
HH member is in farmer organization (1=Yes)	23.658	**0.045
No. adults (>17 yr) literate	2.854	0.631
No. of Tropical Livestock Units	-10.059	0.298
Owns motorcycle (1=Yes)	11.508	0.734
Owns bicycle (1=Yes)	29.001	***0.000
Index of home assets	-9.376	**0.046
Index of productive assets	14.987	0.153
IDA office in village (1=Yes)	45.532	*0.053
Public market in village (1=Yes)	-21.092	0.136
HH in Caala Municipality (1=Yes)	35.726	0.325
HH in Ekunha Municipality (1=Yes)	51.602	0.165
HH in Bailundo Municipality (1=Yes)	-88.211	*0.065
HH in Katchiungo Municipality (1=Yes)	17.482	0.376
HH in Tchicalachuluanga Municipality (1=Yes)	-14.595	0.505
HH in Chiguar Municipality (1=Yes)	-5.821	0.721
HH in Babaera Municipality (1=Yes)	-63.376	**0.048
Distance from village to sede (km)	-0.681	0.295
Road between village and sede in poor condition (1=Yes)	-56.612	***0.001
Seller sought price information prior to sales (1=Yes)	-3.788	0.679
Reported marketing costs (Kw/kg)	-0.526	0.655
Total potato production (kg)	0.570	***0.000

Notes: *, **, *** indicates the corresponding coefficients are significant at the 10%, 5%, and 1% levels, respectively. Coefficients and p-values obtained via bootstrapping at 500 repetitions.

Dependency ratio estimated by dividing No. members <17 yr by household (HH) size. Literacy refers to adults who can read and write. n.a. = not applicable because variable was not included in the regression.

Source: ProRenda survey, Angola, 2009. Estimates weighted to reflect population.

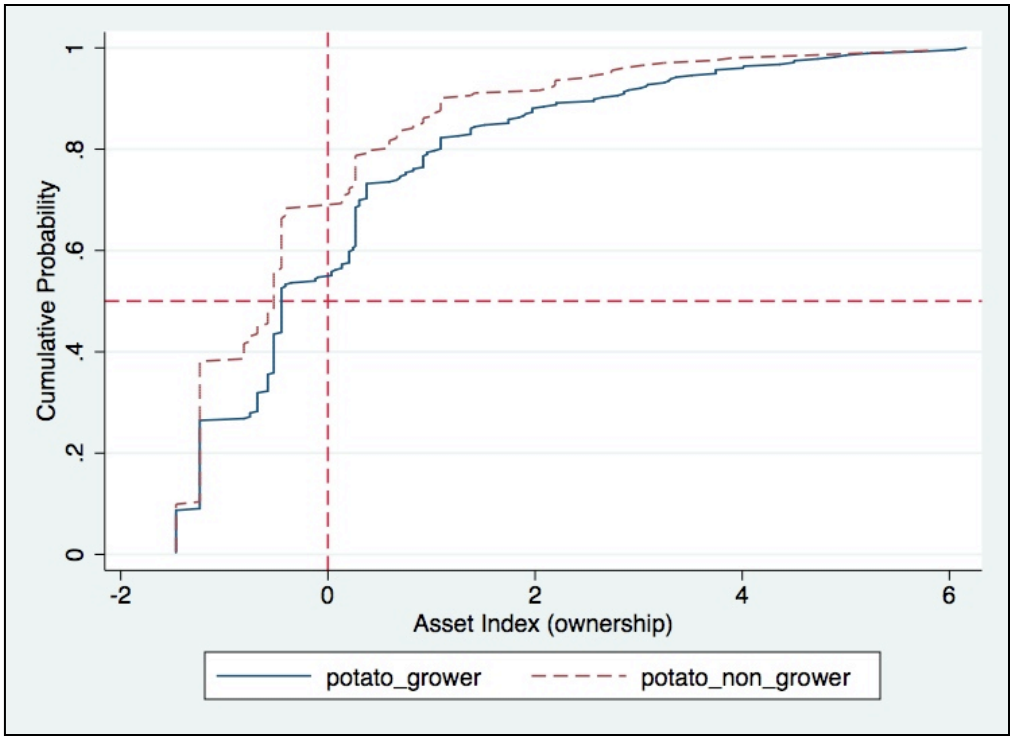

Figure Annex 6.1. Cumulative distribution of asset index by potato growers and non-growers. Central Highlands of Angola, 2009.

REFERENCES

Barrett, C. B. 2008. Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy*, Vol. 33 (4): 299-317.

Bellemare, M. and Barrett, C. B. 2006. An Ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia. *American Journal of Agricultural Economics*, Vol. 88 (2): 324-337.

Benfica, R.; Tschirley, D.; and Boughton, D. 2006. Interlinked Transactions in Cash Cropping Economies: The Determinants of Farmer Participation and Performance in the Zambezi River Valley of Mozambique. Contributed paper presented at the 26th International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006. 15 p.

Boughton, D.; Mather, D.; Barrett, C. B.; Benfica, R.; Abdula, D.; Tschirley, D.; and Cunguara, B. 2007. Market Participation by Rural Households in a Low-Income Country: An Asset-Based Approach Applied to Mozambique. *Faith and Economics*, Vol. 50: 64-101.

Burke, W. J. 2009. Fitting and interpreting Cragg's tobit alternative using Stata. *The Stata Journal*, Vol. 9 (4): 584-592.

Cragg, J. 1971. Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, Vol. 39 (5): 829-844.

De Janvry, A.; Fafchamps, M.; and Sadoulet, E. 1991. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal*, Vol. 101 (409): 1400-1417.

Fafchamps, M. and Vargas-Hill, R. 2005. Selling at the Farmgate or Traveling to Market. *American Journal of Agricultural Economics*, Vol. 87 (3): 717-734.

Filmer, D. and L.H. Pritchett. 2001. Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India. *Demography*, Vol 38 (1): 115-132.

Goetz, S. 1992. A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. *American Journal of Agricultural Economics*, Vol. 74 (2): 444-452.

Gujarati, D. 2003. Basic Econometrics. Fourth Edition. McGraw-Hill Publisher. 1002 p.

Heckman, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, Vol. 47 (1): 153-161.

Key, N.; Sadoulet, E.; and de Janvry, A. 2000. Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics*, Vol. 82 (2): 245-259.

McKenzie, D.J. 2005. Measuring Inequality with Asset Indicators. *Journal of Population Economics*, Vol 18 (2): 229-260.

MINADER and FAO. 2003. Revisão do Sector Agrário e da Estratégia de Segurança Alimentar para Definição de Prioridades de Investimentos (TPC/ANG/2907). Angola, Sistemas de Produção Agrícolas. Documento de Trabalho No. 07. 33 p.

Munslow, B. 1999. Angola: The Politics of Unsustainable Development. *Third World Quarterly*, Vol. 20 (3): 551-568.

Reyes, B.; Donovan, C.; Kelly, V. 2010. Baseline Survey Report. Raising the Incomes of Smallholder Farmers in the Central Highlands of Angola: A Model Project for Improving Agricultural Value Chains in Post-Conflict Nations. ProRenda Project. Report submitted to World Vision. 70 p.

Ricker-Gilbert, J.; Jayne, T.; and Chirwa, E. 2011. Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi. *American Journal of Agricultural Economics.* Vol. 93(1): 26-42.

Sadoulet, E. and de Janvry, A. 1995. Quantitative Development Policy Analysis. The Johns Hopkins University Press; Baltimore and London. 438 p.

StataCorp. 2009. *Stata: Release 11*. Statistical Software. College Station, TX: StataCorp LP.

Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, Vol. 26 (1): 24-36.

Wooldridge, J. M. 2002. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press. 752 p.

Wooldridge, J. M. 2003. Introductory Econometrics: A Modern Approach. 2nd ed. South-Western Publisher. 863 p.

Wooldridge, J. M. 2008. Two-part and hurdle models. EC 821A class notes. Michigan State University. 34 p.

Wooldridge, J. M. 2009. Sample selection and missing data. EC 821B class notes. Michigan State University. 108 p.

World Bank; EC; FAO; UNICEF; UNDP; and WHO. 2007. Angola Public Expenditure Review. Volume II: Sectoral Review. Report No. 39710-AO. 168 p.

World Bank. 2007. Angola: Oil, Broad-based Growth, and Equity. World Bank Country Study Series. Report No. 40531. 156 p.

World Vision. 2008. PRORENDA. Raising the Incomes of Smallholder Farmers in the Central Highlands of Angola: A Model Project for Improving Agricultural Value Chains in Post-Conflict Nations. Grant Proposal. 58 p.